REPORT DOCUMENTATION PAGE		Form Approved OMB NO. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comment regarding this burden estimates or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE March 1997	3. REPORT TYPE AND DATES COVERED Technical	
4. TITLE AND SUBTITLE Constrained Maximum Likelihood Estimators for Superimposed Exponential Signals			5. FUNDING NUMBERS DAAH04-96-1-0082
6. AUTHOR(S) D. Kundu and N. Kannan			
7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES) Center for Multivariate Analysis Department of Statistics 417 Thomas Building Penn State University University Park, PA 16802			8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211			10. SPONSORING / MONITORING AGENCY REPORT NUMBER ARO 35518.10-MA
11. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited.			12 b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

Recently Kundu (1993a) has proposed a non-linear eigenvalue method for finding the maximum likelihood estimators (MLE) of the parameters of undamped exponential signals. It is known to perform better than the previously existing methods like FBLP of Tufts and Kumaresan (1982) or Pisarenko's method (Pisarenko, 1972), in the sense of lower mean squared errors. The solution in general depends on the roots of a polynomial equation. It is observed that the coefficients of the polynomial exhibit a certain symmetry. Since it is known (Crowder, 1984) that the MLE with constraints is more efficient than the unconstrained MLE, modified maximum likelihood method has been suggested to estimate the parameters under these symmetric constraints. It is observed in the simulation study that the mean squared errors of the constrained MLE are closer to the Cramer-Rao lower bound than the ordinary MLE in almost all the situations.

19970321 031

14. SUBJECT TERMS Constrained MLE, Multiple Sinusoids, Prony's Algorithm, EVLP, FBLP		15. NUMBER IF PAGES 36
		16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UL

CONSTRAINED MAXIMUM LIKELIHOOD ESTIMATORS

FOR SUPERIMPOSED EXPONENTIAL SIGNALS

Debasis Kundu and Nandini Kannan

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

CONSTRAINED MAXIMUM LIKELIHOOD ESTIMATORS
FOR SUPERIMPOSED EXPONENTIAL SIGNALS

DEBASIS KUNDU[1,] AND NANDINI KANNAN[2]

[1]Department of Mathematics
Indian Institute of Technology
Kanpur, India

[2]Department of Math., Comp. Sc. and Statistics
University of Texas at San Antonio
Texas-78249, U.S.A.

ABSTRACT

Recently Kundu(1993a) has proposed a non-linear eigenvalue method for finding the maximum likelihood estimators (MLE) of the parameters of undamped exponential signals. It is known to perform better than the previously existing methods like FBLP of Tufts and Kumaresan (1982) or Pisarenko's method (Pisarenko; 1972), in the sense of lower mean squared errors. The solution in general depends on the roots of a polynomial equation. It is observed that the coefficients of the polynomial exhibit a certain symmetry. Since it is known (Crowder; 1984) that the MLE with constraints is more efficient than the unconstrained MLE, modified maximum likelihood method has been suggested to estimate the parameters under these symmetric constraints. It is observed in the simulation study that the mean squared errors of the constrained MLE are closer to the Cramer-Rao lower bound than the ordinary MLE in almost all the situations.

Key Words and Phrases: Constrained MLE, Multiple Sinusoids, Prony's Algorithm, EVLP, FBLP.

1. INTRODUCTION

In signal processing detecting the number of signals and estimating the parameters of the signals are very important problems. In this article, we consider the estimation of the following exponentially undamped sinusoidal signal with additive noise, i.e.

$$Y_n = \sum_{k=1}^{M} \alpha_k e^{i\omega_k n} + \epsilon_n \quad n = 1,\ldots,N \qquad (1.1)$$

Here α_k are unknown complex numbers, $\omega_k \in [0,2\pi)$ are unknown radian frequencies and $i = \sqrt{-1}$. The ϵ_n are independent identically distributed (i.i.d.) complex valued error random variables such that $E(\epsilon_n) = 0$ and $E|\epsilon_n|^2 = \sigma^2$. $\underset{\sim}{\alpha} = (\alpha_1,\ldots,\alpha_M)$, $\underset{\sim}{\omega} = (\omega_1,\ldots,\omega_M)$ are unknown and the ω_k's are assumed to be distinct. The problem is to estimate the unknown parameters, given a sample of size N. Here we assume that M, the number of signals, is known.

Furthermore consider the following non-linear time series regression model;

$$Y_n = \sum_{k=1}^{L} [A_k \cos(\omega_k n) + B_k \sin(\omega_k n)] + \epsilon_n$$

$$n = 1,\ldots,N \quad (1.2)$$

Here A_k and B_k are real amplitudes and ω_k's are frequencies and ϵ_n's are i.i.d. real valued random variables with mean zero and finite variance σ^2. Observe that the model (1.2) can be written as a special case of the model (1.1).

The problem of estimating the parameters of complex sinusoids in noise has received considerable attention in the past several years, (see for example Kay and Marple (1981) and Kay; (1988) Chapter 13) in the

2

Electrical Engineering literature where as a reasonable amount of work has been done starting with the work of Hannan (1971) and Walker (1971),in the Statistics literature also, (see for example Rice and Rosenblatt; (1988) and Kundu (1993b),for estimating real sinusoids. Under the normality assumptions on the error term, the maximum likelihood estimators (MLE) of ω and α are same as the non-linear least squares estimators obtained by minimizing

$$R_N(\underset{\sim}{a},\underset{\sim}{\theta}) = \sum_{n=1}^{N} \left| y_n - \sum_{k=1}^{M} a_k \, e^{in\theta_k} \right|^2 \qquad (1.3)$$

with respect to $\underset{\sim}{\theta} = (\theta_1,\ldots,\theta_M)$ and $\underset{\sim}{a} = (a_1,\ldots,a_M)$. The solution obtained from (1.3) is a consistent estimator of ω and α (Mitra and Kundu; 1993 or Rao and Zhao (1993)). Unfortunately it is known that (Smyth; 1985, Kundu; 1989) the general purpose algorithms such as Gauss Newton, Newton Rapson or their variants often have great difficulty in converging to the optimum solution. Recently Kundu (1993a) gives an efficient iterative procedure to obtain the MLE of ω, the asymptotic stability of the procedure can be found in Kundu (1994).

In estimating the frequencies, the methods like the Pisarenko's method (Pisarenko; 1972), Forward Backward Linear Prediction (FBLP) of Tufts and Kumaresan (1982), EquiVariance Linear Prediction (EVLP) discussed in Bai, Krishnaiah and Zhao (1986) or ·Rao (1988), are based on the classical methods of Prony (1795). The basic MLE Prony method goes back to Osborne (1975). That method was further developed and investigated by Osborne and Smyth (1991, 1994) and Kahn et. al. (1992). Bai, Rao and Chow (1989) refine the

EVLP estimators to produce efficient estimators of the frequencies. The solution in general depends on the roots of a polynomial equation. It is observed that the coefficients of the polynomial equation satisfy some symmetric constraints. Since it is known (Crowder; 1984) that the constrained MLE has lower mean squared error then the unconstrained MLE, we introduce constrained MLE for this problem and compare it with the existing methods using simulation. Observe that once we obtain the non-linear parameters the linear parameters can be obtained using the simple linear regression technique (Kundu, 1994).

The rest of the paper is organized as follows, we provide the Pisarenko's method in Chapter 2. The Modified Prony algorithm is given in Chapter 3, Chapter 4 contains the Constrained maximum likelihood procedure. Some Numerical results are provided in Chapter 5 and we draw conclusions from our results in Chapter 6. In the Appendix we provide some theoretical justification of the convergence of the numerical procedures under the assumptions of independent and identically distributed error random variables.

2. PISARENKO'S METHOD

Prony (1795) suggested a method of solving the non-linear estimation problem. Many standard texts on numerical methods outline this algorithm (Barrodale and Olesky (1981), Froberg (1969)). The algorithm can be extended to the noisy case as follows :

Suppose that the vector $g = (g_0, \ldots, g_M)'$ is such that

$$g_0 + g_1 z + \ldots + g_M z^M = g_M \prod_{j=1}^{M} (z - e^{-i\omega_j}) \qquad (2.1)$$

then for any $n \geq M+1$

$$\sum_{k=0}^{M} g_k Y_{n-k} = \sum_{k=0}^{M} g_k \epsilon_{n-k} \qquad (2.2)$$

where the right hand side of (2.2) is a function of error only. The coefficients g_j are estimated by minimizing

$$\sum_{n=M+1}^{N} \left| \sum_{k=0}^{M} g_k Y_{n-k} \right|^2 \qquad (2.3)$$

subject to the condition $|g| = 1$. Such a method of estimation is known as the Pisarenko's method, (also named as EVLP method by Bai, Krishnaiah and Zhao) (1986).

Now write

$$T = \begin{bmatrix} Y_1 & \cdots\cdots\cdots & Y_{M+1} \\ \\ Y_{N-M} & \cdots\cdots & Y_N \end{bmatrix} \qquad (2.4)$$

and $R = T^*T$, here '*' denotes the complex conjugate transpose of a matrix or of a vector. Therefore the estimator of g can be obtained by minimizing g^*Rg such that $g^*g = 1$. If $\hat{g}$ is the estimator of g so obtained construct the polynomial equation

$$\hat{g}_o + \hat{g}_1 z + \ldots + \hat{g}_M z^M = 0 \qquad (2.5)$$

obtain solutions in the form of

$$\hat{\rho}_1 e^{-i\hat{\omega}_1}, \ldots, \hat{\rho}_M e^{-i\hat{\omega}_M} \quad \text{with } \rho_i > 0$$

$$\text{for } i = 1, 2, \ldots M. \qquad (2.6)$$

and take $\hat{\omega}_1, \ldots, \hat{\omega}_M$ as estimators of $\omega_1, \ldots, \omega_M$. It is

5

shown in Bai, Krishnaiah and Zhao (1986) that $\hat{\omega}$ is a consistent estimate of ω with a convergence rate of $O_p(N^{-1/2})$. It is observed that [Rao and Zhao; (1993) Mitra and Kundu (1993] $\hat{\omega}$ obtained by minimizing (1.2) is a consistent estimator of ω with a convergence rate of $O_p(N^{-3/2})$. On the other hand the FBLP of Tufts and Kumaresan (1982) may not provide consistent estimator of ω (Rao; 1988).

3. MODIFIED PRONY ALGORITHM

A modification to Prony's algorithm have been considered by Marple (1979,1987), using the fact that the roots of the polynomial equation (2.5) should be of unit modulus. Consider the polynomial equation

$$P(z) = g_O + g_1 z + \ldots + g_M z^M = 0 \qquad (3.1)$$

with z_k, $k = 1, \ldots M$ as its roots. Since the z_k's are of unit modulus, we have $z_k^{-1} = \bar{z}_k$. Therefore the polynomial $P(z)$ and

$$R(z) = \bar{g}_M + \bar{g}_{M-1} z + \ldots + \bar{g}_O z^M = 0 \qquad (3.2)$$

have the same roots. Hence

$$\frac{g_k}{\bar{g}_{M-K}} = \frac{g_M}{\bar{g}_O} \qquad\qquad K = 0, \ldots, M$$

or

$$g_K \left[\frac{\bar{g}_O}{g_M} \right]^{1/2} = \bar{g}_{M-K} \left[\frac{g_M}{\bar{g}_O} \right]^{1/2} \qquad K = 0, \ldots M \qquad (3.3)$$

If we define

6

$$b_K = g_K \left[\frac{\overline{g}_o}{g_M} \right]^{1/2} \qquad K = 0,\ldots M \qquad (3.4)$$

then it is easy to see that $b_K = \overline{b}_{M-K}$; $K = 0,\ldots M$. Thus (2.2) can be written as; if y_i satisfies (1.1), then there exists a vector $\underset{\sim}{b} = (b_0, \ldots, b_M)$ such

that $b_k = \overline{b}_{M-k}$; $k = 0,\ldots,M$ and

$$\sum_{k=0}^{M} b_k \, y_{n-k} = \sum_{k=0}^{M} b_k \, \varepsilon_{n-k} \qquad (3.5)$$

for any $n \geq M+1$.

4. CONSTRAINED MLE

Bresler and Macovski (1986) and Kumaresan, Scharf and Shaw (1986) considered the maximum likelihood approach using similar conditions on the coefficients. They transformed the problem to a constrained quadratic minimization problem and solved it iteratively using some standard computer packages. No proof of convergence was provided in these papers. In this section we modify the MLE procedure of Kundu (1993a) using similar conditions on the coefficients. We transform the problem to a non-linear eigenvalue problem and solve it iteratively.

The model (1.1) can be written in the following matrix form;

$$\underset{\sim}{Y} = A(\omega)\underset{\sim}{\alpha} + \underset{\sim}{E} \qquad (4.1)$$

where $\underset{\sim}{Y}^T = (y_1,\ldots,y_N)$, $\underset{\sim}{\alpha}^T = (\alpha_1,\ldots,\alpha_M)$, $\underset{\sim}{\omega}^T = (\omega_1,\ldots,\omega_M)$ and $E^T = (\epsilon_1,\ldots,\epsilon_N)$. $A(\omega)$ is a NxM matrix and its (p,q)th element is $e^{i\omega_q p}$, $p = 1,\ldots,N$, q

7

$= 1,...,M.$

Therefore (1.3) can be written as

$$R_N(\alpha, \omega) = |Y - A(\omega)\alpha|^2 \qquad (4.2)$$

Here the linear parameters α are separable from the non-linear parameters ω. For a fixed value of ω, the minimization of R_N with respect to α is a simple linear regression problem. The solution of $\hat{\alpha}(\omega)$ is

$$\hat{\alpha}(\omega) = (A^*(\omega) A(\omega))^{-1} A^*(\omega)Y \qquad (4.3)$$

Substituting back $\alpha(\omega)$ in (4.2) and denoting $Q(\omega) = R_N(\alpha,\omega)$ we obtain

$$Q(\omega) = Y^*(I - P_A)Y \qquad (4.4)$$

where $P_A = A(A^*A)^{-1}A^*$ is the projection operator in the range space spanned by the columns of the matrix A. Therefore the MLE of ω can be obtained by minimizing (4.4), w.r.t. ω.

Consider the NxN-M matrix X of the following form:

$$X = \begin{bmatrix} \overline{g}_O & 0 & & 0 \\ \vdots & \overline{g}_O & & 0 \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & 0 \\ \overline{g}_M & \vdots & & \overline{g}_O \\ \vdots & \overline{g}_M & & \vdots \\ 0 & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & \overline{g}_M \end{bmatrix} \qquad (4.5)$$

From (2.2) it follows $X^*\mu = 0$, where $\mu' =$

$(Ey_1, \ldots, Ey_N)$. This implies $X^* A(\omega) = 0$, i.e. the

columns of A are orthogonal to the rows of X^* and $I -$

$P_A = P_X = X(X^*X)^{-1}X^*$. Therefore minimizing $Y^*(I-P_A)Y$

with respect to ω is equivalent to minimizing $Y^* P_X Y$

with respect to $g = (g_0, \ldots, g_M)$.

Let $\hat{g}$ be the vector that minimizes

$Y^* X(X^*X)^{-1}X^*Y$. Then the MLE of ω can be obtained

by solving the polynomial equation as described in

(2.5) and (2.6) (see Kundu; (1993a). Similarly the

constrained MLE of ω can be obtained by minimizing

$$Q(g) = Y^* X(X^*X)^{-1}X^*Y \tag{4.6}$$

such that $g^* g = 1$ and $g_k = \bar{g}_{M-k}$; $k = 0, \ldots, M$.

Then the estimator of ω can be obtained by solving the

polynomial equation (2.5). Observe that the roots of

the polynomial equation in the constrained case will be

of the form

$$e^{-i\hat{\omega}_1}, \ldots, e^{-i\hat{\omega}_M} \tag{4.7}$$

instead of (2.6).

The constrained minimization can be done as

follows. The matrix X in the constrained case can be

written as $X = [x_1, \ldots, x_{N-M}]$ where x_i^* is of the form [

$0, g, 0$]. Let U and V denote the real and imaginary

parts of X respectively. Assuming M is odd for

brevity and $g_k = c_k - id_k$ for $k = 0, 1, \ldots, (M-1)/2$.

Then X can be written as

$$X = \sum_{\alpha=0}^{(M-1)/2} (c_\alpha U_\alpha + id_\alpha V_\alpha) \tag{4.8}$$

where U_α, V_α; $\alpha = 0,1,\ldots,\frac{M-1}{2}$ are NxN-M matrices with entries 0 and 1. The minimization of $Q(g) = Y^*X(X^*X)^{-1}X^*Y$ such that $g^*g = 1$ and $g_k = \bar{g}_{M-k}$; $k = 0,\ldots,M$ is equivalent to minimize

$$Q(c, d) + \lambda (c^Tc + d^Td - 1) \tag{4.9}$$

with respect to c, d and λ. Here we denote $c^T = (c_0,\ldots,\frac{c_{M-1}}{2})$, $d^T = (d_0,\ldots,\frac{d_{M-1}}{2})$, $Q(c, d) = Y^*X(X^*X)^{-1}X^*Y$, where X is of the form (4.8) and λ is the Lagrange multiplier. If we differentiate (4.9) with respect to c and d, we obtain a matrix equation of the form

$$D(c^T, d^T) \begin{bmatrix} c \\ d \end{bmatrix} + 2\lambda \begin{bmatrix} c \\ d \end{bmatrix} = 0 \tag{4.10}$$

Here D is a M+1xM+1 matrix. The matrix D can be written in a partitioned form as :

$$D = \begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \tag{4.11}$$

where A,B,C are all $\frac{M+1}{2}$ x $\frac{M+1}{2}$ matrices. We have

$$A_{jk} = Y^*U_j(X^*X)^{-1}U_k^*Y - Y^*X(X^*X)^{-1}(U_j^TU_k + U_k^TU_j)(X^*X)^{-1}X^*Y$$

$$+ Y^*U_k(X^*X)^{-1}U_j^TY \quad j,k = 0,1,\ldots\frac{M-1}{2} \tag{4.12}$$

$$B_{jk} = -iY^*U_j(X^*X)^{-1}V_k^*Y - iY^*X(X^*X)^{-1}(U_j^TV_k - V_k^TU_j)(X^*X)^{-1}X^*Y$$

$$+ iY^*V_k(X^*X)^{-1}U_j^TY \quad j,k = 0,1,\ldots\frac{M-1}{2} \tag{4.13}$$

10

$$C_{jk} = Y^* V_j (X^* X)^{-1} V_k^* Y - Y^* X (X^* X)^{-1} (V_j^T V_k + V_k^T V_j)(X^* X)^{-1} X^* Y$$

$$+ Y^* V_k (X^* X)^{-1} V_j^T Y \qquad j,k = 0,1,\ldots \frac{M-1}{2} \qquad (4.14)$$

Premultiplying the left hand side of (4.10) by (c^T, d^T), we obtain

$$[c^T, d^T] \; D(c^T, d^T) \begin{bmatrix} c \\ \sim \\ d \\ \sim \end{bmatrix} + 2\lambda = 0 \qquad (4.15)$$

It can be easily seen that

$$[c^T, d^T] \; D(c^T, d^T) \begin{bmatrix} c \\ \sim \\ d \\ \sim \end{bmatrix} = 0 \qquad (4.16)$$

which implies $\lambda = 0$. Therefore the constrained minimization of $Q(g)$ is equivalent to solving a matrix equation of the form

$$D(c^T, d^T) \begin{bmatrix} c \\ \sim \\ d \\ \sim \end{bmatrix} = 0 \qquad \text{s.t. } c^T c + d^T d = 1 \qquad (4.17)$$

Similarly when M is even, $c^T = (c_0, \ldots, c_{\frac{M}{2}})$, $d^T = \left(d_0, \ldots, d_{\frac{M}{2}-1}\right)$ and D can be partitioned as

$$D = \begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \qquad (4.18)$$

Here A, B, C are $\frac{M}{2} + 1 \times \frac{M}{2} + 1$, $\frac{M}{2} + 1 \times \frac{M}{2}$, $\frac{M}{2} \times \frac{M}{2}$, matrices respectively. The (j,k)th element of A, B, or C is same as before with appropriate ranges. Note that D is a real symmetric matrix in both the cases, and we need to solve a matrix equation of the form

$$D(x)x = 0 \qquad \text{s.t.} \qquad \|x\| = 1 \qquad (4.19)$$

11

This is a nonlinear eigenvalue problem. $\hat{\underset{\sim}{x}}$ satisfying (4.19) should be an eigenvector corresponding to the zero eigenvalue of the matrix $D(\hat{\underset{\sim}{x}})$. We suggest the following iterative technique similar to that of Osborne (1975) and Kundu (1993a) to solve (4.19) :

$$(D(\underset{\sim}{x}^k) - \lambda^{k+1}I)\underset{\sim}{x}^{k+1} = 0 \qquad \|\underset{\sim}{x}^{k+1}\| = 1 \qquad (4.20)$$

where λ^{k+1} is the eigenvalue closest to zero of $D(\underset{\sim}{x}^k)$ and $\underset{\sim}{x}^{k+1}$ is the corresponding normalized eigenvector. The iterative process should be stopped when λ^{k+1} is small corresponding to $\|D\|$.

The algorithm has the following form :

Step 1: Set an initial value $\underset{\sim}{x}^1$ and normalize it, i.e.

$$\underset{\sim}{x}^1 = \underset{\sim}{x}^1/\|\underset{\sim}{x}^1\| \; ; \qquad i = 1.$$

Step 2: Calculate matrix $D(\underset{\sim}{x}^i)$

Step 3: Find the eigenvalue λ^{i+1} of $D(\underset{\sim}{x}^i)$ closest to zero and normalize the corresponding eigenvector $\underset{\sim}{x}^{i+1}$

Step 4: Test the convergence by checking if

$$|\lambda^{i+1}| < \in \|D\|$$

Step 5: If the test in 4 fails, i:= i+1 and go to Step 2.

12

5. NUMERICAL EXPERIMENTS:

We performed some numerical experiments mainly to compare the performances of the different methods for finite samples. All these simulations are done using the IMSL (1984) random deviate generator. We consider the following two models:

Model 1.

$$Y_n = 2.0 \ e^{1.0in} + 3.0 \ e^{2.0in} + \epsilon_n$$

$$n = 1,\ldots,N. \qquad (5.1)$$

Model 2.

$$y_n = 1.0 \ \cos(2.0n) + 1.0 \ \sin(2.0n) + \epsilon_n \qquad (5.2)$$

For Model 1, ϵ_n are i.i.d. complex valued normal random variables with mean zero and standard deviation σ for both the real and imaginary parts. The real and imaginary parts are taken to be independently distributed. For Model 2 ϵ_n are i.i.d. real valued random variables with mean zero and standard deviation σ. For different values of N and σ, one thousand different data sets were generated. Numerical results are observed for sample sizes N = 20,30,40,50 and σ = .01,.1 and .5. Observe that our method can be applied even if the errors are stationary random variables. For Model 2, we also consider when ϵ_n are moving average error of the form;

$$\epsilon_n = e_n + .5 \ e_{n-1} \qquad (5.3)$$

where e_n's are i.i.d. normal random variables with mean zero and standard deviation σ = .1 and N = 20, 30, 40 and 50.

For each data set, we computed the ordinary MLE (MLE) as described in Kundu (1993a), constrained MLE as in Section 4 (CMLE) and as in Bresler and Macovski

(CMLEBM), the FBLP as in Tufts and Kumaresan (1982) and
the modified EVLP (EVLP) as in Bai, Rao and Chow
(1989). For the MLE and CMLE we use $\varepsilon = 10^{-6}$, the
ordinary EVLP estimator as the starting value and $\|D\|$
to be the largest eigenvalue of the matrix D. For
CMLEBM, we use the same initial value as MLE or CMLE
and 7 steps iterations in all the cases. Observe that
the stopping rule of MLE or CMLE is quite different
than that of CMLEBM. It is observed that for MLE or
CMLE the iteration converged in 7 steps in all the
cases for both the models, with the above stopping
criterion. Therefore to have a reasonable comparison
of MLE or CMLE with CMLEBM we use 7 steps iterations
in all the cases in both the models for CMLEBM also.
The mean estimates of the frequencies and their average
mean squared errors (MSE) over one thousand
replications are reported.

Observe that we write (5.2) in the form of (1.1)
with $\omega_1 = 2.0$ and $\omega_2 = -2.0$. We estimated ω_1 and ω_2 by
the different methods as described above. In all our
simulations for all the methods the two roots of (2.5)
came in the complex conjugate form, we obtained the
estimator of ω from that. We also obtain the
estimators of the linear parameters from (4.3) in all
the cases.

We reported all the results in Table 1-4 for Model
1 and in Table 5 for Model 2 when the errors are i.i.d
and when the errors are from (5.3), the results are
reported in Table 6. In each table we also put the
Cramer-Rao lower bounds (CRLB) for comparison.

14

Table 1a
Model 1
$N = 20, \ \sigma = .01$

Method \	ω_1	ω_2	α_1	α_2
PARAMETERS	1.00000	2.00000	2.00000	3.00000
(CRLB)	(.194E-3)	(.129E-3)	(.224E-2)	(.224E-2)
MLE	1.00007	1.99989	2.00025	3.00010
	(.275E-3)	(.201E-3)	(.246E-2)	(.247E-2)
CMLE	1.00000	1.99997	2.00012	3.00005
	(.215E-3)	(.146E-3)	(.235E-2)	(.242E-2)
CMLEBM	1.00001	1.99997	2.00010	3.00005
	(.223E-3)	(.152E-3)	(.240E-2)	(.242E-2)
FBLP	1.00014	2.00004	3.99996	5.50002
	(.560E-3)	(.220E-3)	(.287E-2)	(.277E-2)
EVLP	0.99982	2.00004	2.00011	3.00040
	(.586E-3)	(.217E-3)	(.304E-2)	(.280E-2)

Table 1b
$N = 20, \ \sigma = .1$

Method \	ω_1	ω_2	α_1	α_2
PARAMETERS	1.00000	2.00000	2.00000	3.00000
(CRLB)	(.194E-2)	(.129E-2)	(.224E-1)	(.224E-1)
MLE	1.00009	1.99983	2.00044	3.00007
	(.280E-2)	(.197E-2)	(.240E-1)	(.251E-1)
CMLE	1.00000	1.99992	2.00010	3.00005
	(.217E-2)	(.149E-2)	(.230E-1)	(.244E-1)
CMLEBM	1.00001	1.99991	2.00010	3.00013
	(.225E-2)	(.149E-2)	(.230E-1)	(.247E-1)
FBLP	2.00001	3.49990	3.99996	5.50002
	(.531E-2)	(.230E-2)	(.311E-1)	(.288E-1)
EVLP	2.00007	3.49996	3.99995	5.50001
	(.521E-2)	(.225E-2)	(.312E-1)	(.282E-1)

15

Table 1c
N = 20, σ = .5

Method \	ω_1	ω_2	α_1	α_2
PARAMETERS	1.00000	2.00000	2.00000	3.00000
(CRLB)	(.968E-2)	(.645E-2)	(.112E-0)	(.112E-0)
MLE	1.00007	1.99989	2.00025	3.00010
	(.176E-1)	(.991E-2)	(.201E-0)	(.182E-0)
CMLE	1.00029	1.99906	1.98828	2.99803
	(.108E-1)	(.726E-2)	(.123E-0)	(.119E-0)
CMLEBM	1.00054	1.99891	1.98650	2.99690
	(.113E-1)	(.754E-2)	(.132E-0)	(.118E-0)
FBLP	0.98911	2.01315	1.85643	2.88781
	(.391E-1)	(.215E-1)	(.340E-0)	(.261E-0)
EVLP	0.98619	1.99715	1.76575	2.88092
	(.398E-1)	(.229E-1)	(.324E-0)	(.254E-0)

Table 2a
Model 1
N = 30, σ = .01

Method \	ω_1	ω_2	α_1	α_2
PARAMETERS	1.00000	2.00000	2.00000	3.00000
(CRLB)	(.105E-3)	(.703E-4)	(.182E-2)	(.182E-2)
MLE	1.00007	1.99989	2.00025	3.00010
	(.201E-3)	(.111E-3)	(.202E-2)	(.219E-2)
CMLE	1.00001	1.99999	2.00016	2.99916
	(.140E-3)	(.712E-4)	(.193E-2)	(.202E-2)
CMLEBM	1.00001	1.99998	2.00014	2.99995
	(.146E-3)	(.739E-4)	(.193E-2)	(.201E-2)
FBLP	0.99983	2.00014	1.99996	3.00190
	(.299E-3)	(.250E-3)	(.270E-2)	(.251E-2)
EVLP	0.99984	1.99978	2.00011	3.00040
	(.316E-3)	(.276E-3)	(.281E-2)	(.277E-2)

Table 2b
$N = 30, \; \sigma = .1$

Method \	ω_1	ω_2	α_1	α_2
PARAMETERS	1.00000	2.00000	2.00000	3.00000
(CRLB)	(.105E-2)	(.703E-3)	(.182E-1)	(.182E-1)
MLE	1.00021	2.00095	2.00011	2.99981
	(.201E-2)	(.991E-3)	(.231E-1)	(.243E-1)
CMLE	0.99989	1.99989	2.00009	2.99891
	(.139E-2)	(.722E-3)	(.187E-1)	(.233E-1)
CMLEBM	1.00014	1.99982	2.00009	2.99959
	(.145E-2)	(.741E-3)	(.191E-1)	(.211E-1)
FBLP	1.00011	2.00019	1.99563	2.99856
	(.293E-2)	(.199E-2)	(.281E-1)	(.279E-1)
EVLP	0.99983	1.99776	2.01350	2.98708
	(.310E-2)	(.195E-2)	(.291E-1)	(.277E-1)

Table 2c
$N = 30, \; \sigma = .5$

Method \	ω_1	ω_2	α_1	α_2
PARAMETERS	1.00000	2.00000	2.00000	3.00000
(CRLB)	(.527E-2)	(.351E-2)	(.913E-1)	(.913E-1)
MLE	0.99987	2.00012	2.00011	3.00121
	(.880E-2)	(.481E-2)	(.991E-1)	(.972E-1)
CMLE	1.00091	1.99909	2.04309	2.99037
	(.739E-2)	(.412E-2)	(.920E-1)	(.921E-1)
CMLEBM	1.00094	1.99913	2.04420	2.99817
	(.743E-2)	(.416E-2)	(.928E-1)	(.923E-1)
FBLP	0.98763	1.99763	2.01341	2.97865
	(.268E-1)	(.175E-1)	(.299E-0)	(.239E-0)
EVLP	0.98866	1.98847	1.81846	2.78936
	(.276E-1)	(.187E-1)	(.306E-0)	(.249E-0)

17

Table 3a
Model 1
$N = 40$, $\sigma = .01$

Method	ω_1	ω_2	α_1	α_2
PARAMETERS	1.00000	2.00000	2.00000	3.00000
(CRLB)	(.685E-4)	(.456E-4)	(.158E-2)	(.158E-2)
MLE	1.00003	1.99948	2.00012	2.99919
	(.891E-4)	(.591E-4)	(.189E-1)	(.211E-1)
CMLE	1.00001	1.99998	2.00091	2.99960
	(.793E-4)	(.495E-4)	(.176E-1)	(.173E-1)
CMLEBM	1.00001	1.99998	2.00019	2.99987
	(.806E-4)	(.505E-4)	(.176E-1)	(.179E-1)
FBLP	1.00017	2.00011	1.99956	2.99281
	(.248E-3)	(.199E-3)	(.199E-2)	(.210E-2)
EVLP	0.99899	1.99981	1.99962	2.99981
	(.256E-3)	(.196E-3)	(.201E-2)	(.202E-2)

Table 3b

$N = 40$, $\sigma = .1$

Method	ω_1	ω_2	α_1	α_2
PARAMETERS	1.00000	2.00000	2.00000	3.00000
(CRLB)	(.685E-3)	(.456E-3)	(.158E-1)	(.158E-1)
MLE	0.99979	2.00091	1.99911	2.99891
	(.911E-3)	(.685E-3)	(.210E-1)	(.206E-1)
CMLE	1.00012	1.99981	2.00889	2.99698
	(.793E-3)	(.496E-3)	(.175E-1)	(.163E-1)
CMLEBM	1.00014	1.99987	2.00886	2.99789
	(.801E-3)	(.499E-3)	(.181E-1)	(.176E-1)
FBLP	0.99923	2.00178	1.99203	3.00879
	(.269E-2)	(.888E-3)	(.271E-1)	(.259E-1)
EVLP	1.00076	1.99786	2.00171	3.00932
	(.275E-2)	(.885E-3)	(.275E-1)	(.261E-1)

18

Table 3c
N = 40, σ = .5

Method \	ω_1	ω_2	α_1	α_2
PARAMETERS	1.00000	2.00000	2.00000	3.00000
(CRLB)	(.342E-2)	(.228E-2)	(.790E-1)	(.790E-1)
MLE	0.99764	2.00120	1.99910	3.00019
	(.423E-2)	(.298E-2)	(.899E-1)	(.911E-1)
CMLE	1.00083	1.99920	2.03879	2.98082
	(.395E-2)	(.248E-2)	(.842E-1)	(.817E-1)
CMLEBM	1.00089	2.00011	2.03563	2.99914
	(.396E-2)	(.255E-2)	(.851E-1)	(.815E-1)
FBLP	1.09805	2.00128	2.02451	3.01395
	(.218E-1)	(.936E-2)	(.311E-0)	(.231E-0)
EVLP	1.00165	1.99765	1.91876	2.88976
	(.211E-1)	(.989E-2)	(.299E-0)	(.241E-0)

Table 4a
Model 1
N = 50, σ = .01

Method \	ω_1	ω_2	α_1	α_2
PARAMETERS	1.00000	2.00000	2.00000	3.00000
(CRLB)	(.489E-4)	(.326E-4)	(.141E-2)	(.141E-2)
MLE	0.99998	1.99988	2.00009	2.99999
	(.601E-4)	(.380E-4)	(.181E-2)	(.178E-2)
CMLE	1.00000	1.99999	1.99998	2.99991
	(.570E-4)	(.320E-4)	(.171E-2)	(.140E-2)
CMLEBM	0.99999	1.99999	2.00001	2.99990
	(.575E-4)	(.326E-4)	(.176E-2)	(.141E-2)
FBLP	1.00011	1.99920	1.99931	3.00016
	(.158E-3)	(.139E-3)	(.189E-2)	(.189E-2)
EVLP	0.99994	2.00010	2.00078	3.00001
	(.159E-3)	(.136E-3)	(.198E-2)	(.198E-2)

Table 4b
N = 50, σ = .1

Method \	ω_1	ω_2	α_1	α_2
PARAMETERS (CRLB)	1.00000 (.489E-3)	2.00000 (.326E-3)	2.00000 (.141E-1)	3.00000 (.141E-1)
MLE	1.00012 (.710E-3)	2.00088 (.389E-3)	1.99962 (.191E-1)	3.00019 (.199E-1)
CMLE	0.99992 (.670E-3)	1.99993 (.335E-3)	2.00769 (.167E-1)	2.99905 (.159E-1)
CMLEBM	1.00056 (.675E-3)	2.00091 (.340E-3)	2.00239 (.170E-1)	3.00317 (.169E-1)
FBLP	0.99918 (.238E-2)	2.00098 (.790E-3)	2.00270 (.262E-1)	3.00349 (.244E-1)
EVLP	1.00072 (.244E-2)	1.99911 (.787E-3)	2.00252 (.251E-1)	3.00013 (.249E-1)

Table 4c
N = 50, σ = .5

Method \	ω_1	ω_2	α_1	α_2
PARAMETERS (CRLB)	1.00000 (.245E-2)	2.00000 (.163E-2)	2.00000 (.707E-1)	3.00000 (.707E-1)
MLE	1.00569 (.375E-2)	2.00253 (.218E-2)	1.99987 (.849E-1)	2.99911 (.811E-1)
CMLE	0.99971 (.327E-2)	1.99969 (.198E-2)	2.03347 (.830E-1)	2.99546 (.710E-1)
CMLEBM	0.99976 (.335E-2)	2.00021 (.205E-2)	1.99281 (.833E-1)	2.99786 (.707E-1)
FBLP	1.00198 (.190E-1)	1.99875 (.788E-2)	2.00320 (.303E-0)	3.00237 (.225E-0)
EVLP	1.00880 (.198E-1)	2.00210 (.779E-2)	1.91642 (.288E-0)	2.95185 (.239E-0)

Table 5a
Model 2
$\sigma = .01, \omega = 2.0$

Sample Size\ Methods	20	30	40	50
CRLB	2.00000 (.387E-3)	2.00000 (.211E-3)	2.00000 (.137E-3)	2.00000 (.980E-4)
MLE	1.99989 (.401E-3)	1.99996 (.311E-3)	1.99995 (.201E-3)	1.99989 (.123E-3)
CMLE	1.99996 (.390E-3)	1.99997 (.219E-3)	1.99998 (.142E-3)	1.99999 (.988E-4)
CMLEBM	1.99998 (.390E-3)	1.99999 (.220E-3)	1.99998 (.144E-3)	1.99998 (.995E-4)
FBLP	1.99987 (.498E-3)	1.99990 (.437E-3)	1.99989 (.289E-3)	1.99995 (.241E-3)
EVLP	1.99988 (.500E-3)	1.99989 (.431E-3)	1.99999 (.299E-3)	1.99998 (.238E-3)

$\sigma = .1, \omega = 2.0$

Sample Size\ Methods	20	30	40	50
CRLB	2.00000 (.387E-2)	2.00000 (.211E-2)	2.00000 (.137E-2)	2.00000 (.980E-3)
MLE	1.99983 (.403E-2)	1.99989 (.314E-2)	1.99991 (.199E-2)	2.00001 (.119E-2)
CMLE	1.99993 (.388E-2)	2.00061 (.220E-2)	1.99991 (.141E-2)	2.00003 (.990E-3)
CMLEBM	2.00017 (.388E-2)	1.99992 (.224E-2)	1.99999 (.141E-2)	1.99992 (.994E-3)
FBLP	2.00011 (.501E-2)	2.00032 (.441E-2)	1.99994 (.286E-2)	1.99988 (.244E-2)
EVLP	2.00023 (.497E-2)	1.99965 (.434E-2)	1.99989 (.297E-2)	1.99988 (.240E-3)

21

$\sigma = .5, \omega = 2.0$

Sample Size\ Methods	20	30	40	50
CRLB	2.00000 (.194E-1)	2.00000 (.105E-1)	2.00000 (.685E-2)	2.00000 (.490E-2)
MLE	2.00011 (.225E-1)	2.00004 (.218E-1)	1.99995 (.789E-2)	1.99997 (.588E-2)
CMLE	2.00071 (.196E-1)	1.99918 (.109E-1)	1.99954 (.699E-2)	2.00013 (.498E-2)
CMLEBM	1.99901 (.198E-1)	2.00018 (.110E-1)	2.00001 (.705E-2)	2.00071 (.500E-2)
FBLP	2.00088 (.401E-1)	2.00098 (.399E-1)	1.99934 (.303E-1)	2.00108 (.255E-1)
EVLP	1.99975 (.402E-1)	1.99955 (.388E-1)	2.00152 (.298E-1)	1.99917 (.266E-1)

Table 5b
Model 2
$\sigma = .01, A = 1.0$

Sample Size\ Methods	20	30	40	50
CRLB	1.00000 (.500E-2)	1.00000 (.408E-2)	1.00000 (.354E-2)	1.00000 (.316E-2)
MLE	1.00765 (.559E-2)	0.99696 (.501E-2)	1.00145 (.488E-2)	1.00132 (.370E-2)
CMLE	1.00249 (.556E-2)	1.00201 (.457E-2)	1.00158 (.400E-2)	0.99891 (.339E-2)
CMLEBM	0.96269 (.573E-2)	1.00709 (.458E-2)	0.99290 (.428E-2)	0.99852 (.343E-2)
FBLP	1.00987 (.858E-2)	0.99189 (.757E-2)	1.00179 (.689E-2)	1.00012 (.521E-2)
EVLP	1.00266 (.846E-2)	1.00196 (.762E-2)	0.99920 (.687E-2)	1.00236 (.538E-2)

$$\sigma = .1, \ A = 1.0$$

Sample Size\Methods	20	30	40	50
CRLB	1.00000 (.500E-1)	1.00000 (.408E-1)	1.00000 (.354E-1)	1.00000 (.316E-1)
MLE	0.98701 (.588E-1)	0.97236 (.510E-1)	1.00561 (.491E-1)	1.01143 (.390E-1)
CMLE	1.01661 (.561E-1)	1.00728 (.453E-1)	1.01239 (.409E-1)	1.01226 (.344E-1)
CMLEBM	1.06559 (.568E-1)	1.02134 (.455E-1)	1.01187 (.411E-1)	1.00333 (.341E-1)
FBLP	1.02129 (.784E-1)	1.03458 (.771E-1)	1.00977 (.680E-1)	1.01949 (.515E-1)
EVLP	0.98591 (.800E-1)	0.97291 (.751E-1)	0.99350 (.679E-1)	1.00196 (.511E-1)

$$\sigma = .5, \ A = 1.0$$

Sample Size\Methods	20	30	40	50
CRLB	1.00000 (.250E-0)	1.00000 (.204E-0)	1.00000 (.177E-0)	1.00000 (.158E-1)
MLE	1.08614 (.318E-0)	0.94289 (.287E-0)	0.95671 (.255E-0)	1.07197 (.243E-0)
CMLE	1.06013 (.292E-0)	1.01013 (.218E-0)	1.01601 (.201E-0)	1.01038 (.169E-0)
CMLEBM	1.01685 (.309E-0)	0.97094 (.234E-0)	0.99316 (.214E-0)	0.99043 (.177E-0)
FBLP	0.89376 (.558E-0)	1.21643 (.534E-0)	0.90185 (.499E-0)	1.10967 (.320E-0)
EVLP	0.88750 (.562E-0)	0.89921 (.501E-0)	0.89958 (.468E-0)	1.00196 (.355E-0)

Table 6a

Model 2
Moving Average Error
$\sigma = .1, \omega = 2.0$

Sample Size\ Methods	20	30	40	50
CRLB	2.00000 (.353E-2)	2.00000 (.193E-2)	2.00000 (.125E-2)	2.00000 (.895E-3)
MLE	1.99998 (.399E-2)	2.00008 (.279E-2)	1.99999 (.177E-2)	2.00006 (.171E-2)
CMLE	2.00015 (.372E-2)	2.00009 (.201E-2)	2.00001 (.132E-2)	1.99999 (.901E-3)
CMLEBM	2.00018 (.381E-2)	2.00000 (.203E-2)	2.00006 (.140E-2)	2.00001 (.911E-2)
FBLP	1.99619 (.757E-2)	1.99658 (.652E-2)	1.99730 (.427E-2)	1.99641 (.299E-2)
EVLP	2.00610 (.779E-2)	2.00186 (.645E-2)	1.99851 (.441E-2)	2.00181 (.315E-2)

Table 6b
$\sigma = .1, A = 1.0$

Sample Size\ Methods	20	30	40	50
CRLB	1.00000 (.460E-1)	1.00000 (.372E-1)	1.00000 (.333E-1)	1.00000 (.289E-1)
MLE	0.96120 (.571E-1)	1.00961 (.511E-1)	0.99280 (.451E-1)	0.99549 (.377E-1)
CMLE	0.99256 (.501E-1)	1.00113 (.409E-1)	1.00147 (.367E-1)	0.99990 (.291E-1)
CMLEBM	1.00825 (.517E-1)	1.01048 (.411E-1)	1.00290 (.375E-1)	1.00034 (.301E-1)
FBLP	1.03703 (.827E-1)	1.04686 (.811E-1)	1.05334 (.723E-1)	0.93356 (.700E-1)
EVLP	0.94707 (.901E-1)	0.93546 (.859E-1)	1.08556 (.801E-1)	0.91957 (.733E-1)

24

6. CONCLUSIONS:

In this paper we try to estimate the parameters of a complex sinusoidal signal. Since any algorithm analyzing a complex signal can be used easily for analyzing the corresponding real signal (Kumaresan and Tufts; 1982) therefore our algorithm can also be used easily to analyze the corresponding real signal. It is well known that the estimators depend on the root of a certain polynomial equation. It is observed that the coefficients of the polynomial exhibit a certain symmetry. In this paper we modify the maximum likelihood method of Kundu (1993a) by using this symmetry constraints. Bresler and Macovski (1986) and Kumaresan, Scharf and Shaw(1986) also obtained the maximum likelihood estimators using this symmetry. Their methods are very much similar to each other (see Bresler and Macovski; 1986). Kumaresan, Scharf and Shaw (1986) suggested to use some standard constrained optimization packages whereas Bresler and Macovski (1986) obtained the explicit normal equations to solve this constrained problem. Since both the methods are almost same, we report the results of Bresler and Macovski (1986) only.

The numerical results confirm the satisfactory performance of the CMLE algorithm. The mean estimators over one thousand replications show that the sinusoidal frequencies can be estimated unbiasedly by all the methods. But as far as MSE is concerned, the CMLE has the lowest MSE in almost all the cases. The MSE of the CMLE reaches the CRLB in some situations. As N increases or σ decreases, the MSE of the estimators decreases. When N is large and σ is small, MLE, CMLE and CMLEBM behave quite similarly but for small N and

large value σ CMLE behaves marginally better than CMLEBM. It is observed that using the same number of iterations, the CMLE has a clear advantage over CMLEBM in terms of MSE.

Now to compare the computational merit of different iterative procedures, observe that the main computation involves in each iteration in inverting a N-M x N-M matrix X^*X, where X is of the form (4.5). Since X^*X is a banded Toeplitz matrix with bandwidth 2M+1 the inverse computation can be done very efficiently (see Kumar;1985 or Kumaresan, Scharf and Shaw (1986)). In each iteration to calculate CMLEBM it is required to solve a set of M+1 linear equations of M+1 unknowns and to compute CMLE it is required to obtain the eigenvalues and eigenvectors of a real symmetric matrix of order M+1 x M+1. Observe that in MLE computation we need to calculate the eigenvalues and eigenvectors of a Hermitian matrix of order M+1xM+1. So CMLE involves less computation than MLE and CMLEBM involves less amount of computation than CMLE in each iteration. But usually M is very small compared to N in practice, so computationally it does not make too much of a difference in each iteration calculation.

Another important problem needs discussion is the choice of initial value for the iterative procedure. It is well known that in this problem unless the starting values are reasonably good it is very likely that any iterative procedure converges to a local minimum rather than a global minimum. The work of Rice and Rosenblatt (1988) showed that there are local minimum of the sums of squares with respect to the $\underset{\sim}{\omega}$ which are O(1/N) apart. Since the EVLP method gives a convergence rate of only $O(1/\sqrt{N})$, one would not expect it to be good

26

enough as a starting value to find the global minimum to the sums of squares. Although we observed in our simulation study that the EVLP estimator is a reasonable starting point. Bresler and Macovski (1986) and Kumaresan, Scharf and Shaw (1986) also used the same starting point and obtained global convergence in their simulations. We observe that FBLP also can be used as a reasonable starting value. We perform the experiments with both the starting values and it is observed from any one of them all the algorithm namely MLE, CMLE or CMLEBM converge to the same global minimum. Only difference it makes in the number of iteration and that is also very marginal. It seems if σ is not too large, then EVLP or FBLP estimator can be used as a initial guess to obtain the global optimum. For large σ, it is not very easy to choose a good starting value without having any prior information about the parameters. It seems more work is needed in this direction.

Observe that the symmetry constraint of the polynomial coefficients obtained in Section 3, is necessary but not sufficient for the polynomial roots to be on the unit circle (see Kay and Marple; 1981). The constrained algorithm may return real roots that are reciprocal of one another. But in a neighborhood of the true value, the symmetry constraint is sufficient also. We did not encounter this problem in our simulation study in any model.

Recently Kahn et al (1992) has studied extensively the effect of different scaling factor such as $|g| = 1$ or $g_o = 1$ etc. on the consistency of Prony's method. It is observed that different scaling factors play an important role in the asymptotic inference. It can be exploited here also. It seems extensive study of

27

different scaling factors is needed for this problem.

Observe that the MLE, CMLE and CMLEBM are asymptotically equivalent. It is known (Mitra and Kundu; 1993, Rao and Zhao; 1993) that the asymptotic variance of the maximum likelihood estimates are equal to the CRLB. It is interesting to note that the variances for the least squares estimates are rather greater than the corresponding CRLB in some cases. Since it is a highly non-linear problem, we believe it is due to the poor asymptotic approximation.

Rice and Rosenblatt (1988) observed that unless the frequencies are estimated with the precision $O(N^{-1})$, the corresponding estimators of the linear parameters, say by (4.3), may not be consistent. So although MLE, CMLE or CMLEBM are computationally more expensive than FBLP or EVLP, it may be advisable to use them to obtain consistent estimates of both linear as well as non linear parameters. Among the MLE, CMLE and CMLEBM we recommend to use CMLE one since it has the lowest mean squared errors.

APPENDIX

In this Appendix we try to give some theoretical support for the convergence of the iterative procedure described in (4.20), particularly when the errors are independent and identically distributed. To prove the asymptotic stability of the iterative process (4.20), we need to make the assumption that the true parameter value $(\underset{\sim}{\alpha}, \underset{\sim}{\omega}) = (\alpha_1, \ldots, \alpha_M, \omega_1, \ldots, \omega_M)$ is an interior point of the parameter space. We take M is odd for brevity. The proof for M even is exactly the same. The iterative procedure (4.20) can be expressed as

$$F(c^k, d^k) = (c^{k+1}, d^{k+1}) \qquad (A.1)$$

where c^k and d^k are the first and the last $(M+1)/2$ elements of the vector x^k, i.e. $x^k = (c^k, d^k)$, where $c^k = (c_0^k, \ldots, c_{(M-1)/2}^k)$ and $d^k = (d_0^k, \ldots, d_{(M-1)/2}^k)$. Here the function $F: \mathbb{R}^{M+1} \to \mathbb{R}^{M+1}$ is defined implicitly through (4.20). Therefore the convergence matrix $\dot{F}$ of the iterative process can be written as

$$
\dot{F}(\hat{c}, \hat{d}) = \begin{bmatrix} \dfrac{dc^{k+1}}{dc^k} & \dfrac{dd^{k+1}}{dc^k} \\[2ex] \dfrac{dc^{k+1}}{dd^k} & \dfrac{dd^{k+1}}{dd^k} \end{bmatrix}_{(c,d)=(\hat{c},\hat{d})} \tag{A.2}
$$

Here $\hat{x} = (\hat{c}, \hat{d})$ is the solution of (4.19) and all the four sub matrices are of the order $(M+1)/2 \times (M+1)/2$. The (i,j) th element of dc^{k+1}/dc^k is obtained by taking the partial derivative of c_i^{k+1} with respect to c_j^k and the other elements are also similarly defined. Finally the matrix $\dot{F}(\hat{c}, \hat{d})$ is obtained by evaluating all the elements at the point $\hat{x} = (\hat{c}, \hat{d})$. The sufficient condition that $\hat{x}$ is a point of attraction of the iterative process (4.20) is that the spectral density of $\dot{F}(\hat{c}, \hat{d})$ is less than one (Ortega and Rheinboldt, 1970). Writing (4.20) explicitly using the expressions of D from (4.11) to (4.14), we obtain

$$
\sum_{j=0}^{(M-1)/2} A_{ij} c_j^{k+1} + \sum_{j=0}^{(M-1)/2} B_{ij} d_j^{k+1} = \lambda^{(k+1)} c_i^{k+1} \tag{A.3a}
$$

$$
\sum_{j=0}^{(M-1)/2} B_{ij}^* c_j^{k+1} + \sum_{j=0}^{(M-1)/2} C_{ij} d_j^{k+1} = \lambda^{(k+1)} d_i^{k+1} \tag{A.3b}
$$

$i = 0, 1, \ldots, (M-1)/2$ and $\|(c^{k+1}, d^{k+1})\| = 1$.
Here all the matrix A, B and C are evaluated at $x^k =$

(c^k, d^k). The elements of the matrix $\dot{F}(\hat{c}, \hat{d})$ can be obtained by differentiating both sides of (A.3a) and (A.3b) with respect to c_l^k for $l = 0, 1, \ldots, (M-1)/2$ and evaluating at the point $\hat{x}$ yields

$$A(\hat{c}, \hat{d}) \frac{d\hat{c}^T}{dc_l} + B(\hat{c}, \hat{d}) \frac{d\hat{d}^T}{dc_l} + R_l(\hat{c}, \hat{d}) =$$

$$= \frac{d\hat{\lambda}}{dc_l} \hat{c}^T + \hat{\lambda} \frac{d\hat{c}^T}{dc_l} \qquad (A.4a)$$

$$B^*(\hat{c}, \hat{d}) \frac{d\hat{c}^T}{dc_l} + C(\hat{c}, \hat{d}) \frac{d\hat{d}^T}{dc_l} + S_l(\hat{c}, \hat{d}) =$$

$$= \frac{d\hat{\lambda}}{dd_l} \hat{d}^T + \hat{\lambda} \frac{d\hat{d}^T}{dc_l} \qquad (A.4a)$$

where $R_l(\hat{c}, \hat{d})$ is an $(M+1)/2$ vector whose jth. element for $j = 0, 1, \ldots, (M-1)/2$ is given by

$$
\begin{aligned}
R_{lj}(\hat{c}, \hat{d}) = \\
& - Y^* U_j (X^* X)^{-1} (U_l^T X + X^* U_l)(X^* X)^{-1} X^* Y \\
& - Y^* U_l (X^* X)^{-1} (U_j^T X + X^* U_j)(X^* X)^{-1} X^* Y \\
& + Y^* X (X^* X)^{-1} (U_l^T X + X^* U_l)(X^* X)^{-1} (U_l^T X \\
& + X^* U_j)(X^* X)^{-1} X^* Y + Y^* X (X^* X)^{-1} (U_j^T X \\
& + X^* U_j)(X^* X)^{-1} (U_l^T X + X^* U_l)(X^* X)^{-1} X^* Y \\
& - Y^* X (X^* X)^{-1} (U_j^T X + X^* U_j)(X^* X)^{-1} (U_l^T Y \\
& - Y^* X (X^* X)^{-1} (U_l^T X + X^* U_l)(X^* X)^{-1} (U_j^T Y \qquad (A.5)
\end{aligned}
$$

and the $(M+1)/2$ vector $S_l(\hat{c}, \hat{d})$ is also similarly defined. Here the ith. element of $d\hat{c}^T/dc_l$ for $i = 0, \ldots, (M-1)/2$ is obtained by taking the partial derivative of c_i^{k+1} with respect to c_l^k and evaluating it at $(\hat{c}, \hat{d})$. $d\hat{d}^T/dc_l$ and $d\hat{\lambda}^T/dc_l$ are also similarly

defined.

It is easy to see that because all the quantities inside the parenthesis in

$$\lim_{n \longrightarrow \infty} \left\{ \frac{d\hat{\lambda}}{dc_l} \hat{c}^T + \hat{\lambda} \frac{d\hat{c}^T}{dc_l} \right\} = 0$$

for $l = 0, 1, \ldots, (M-1)/2$ (A.6a)

and $$\lim_{n \longrightarrow \infty} \left\{ \frac{d\hat{\lambda}}{dc_l} \hat{d}^T + \hat{\lambda} \frac{d\hat{d}^T}{dc_l} \right\} = 0$$

for $l = 0, 1, \ldots, (M-1)/2$ (A.6b)

both (A.6a) and (A.6b) are bounded over a compact set. Therefore from (A.4) we obtain

$$\lim_{n \to \infty} \left\{ \frac{1}{n} \begin{bmatrix} A(\hat{c},\hat{d}) & B(\hat{c},\hat{d}) \\ B^*(\hat{c},\hat{d}) & C(\hat{c},\hat{d}) \end{bmatrix} \begin{bmatrix} \frac{d\hat{c}^T}{dc_l} \\ \frac{d\hat{d}^T}{dc_l} \end{bmatrix} + \frac{1}{n} \begin{bmatrix} R(\hat{c},\hat{d}) \\ S(\hat{c},\hat{d}) \end{bmatrix} \right\}$$

$= 0$ (A.7)

The main idea of the proof of the asymptotic stability of the iterative process is as follows; $(1/n)R_l(\hat{c},\hat{d})$ and $(1/n)S_l(\hat{c},\hat{d})$ converges to zero vectors almost surely for $l = 0, 1, \ldots, (M-1)/2$ and $(1/n)D(\hat{c},\hat{d})$, where

$$D(\hat{c},\hat{d}) = \begin{bmatrix} A(\hat{c},\hat{d}) & B(\hat{c},\hat{d}) \\ B^*(\hat{c},\hat{d}) & C(\hat{c},\hat{d}) \end{bmatrix}$$ (A.8)

converges to a positive semi definite matrix with null space spanned by (c^o, d^o), the true value of (c,d), corresponding to the true parameter value ω. Since
$$c\hat{c}^T + d\hat{d}^T = 1$$ (A.9)
therefore

$$\hat{c}\,\frac{d\hat{c}^{\mathbf{T}}}{dc_{\iota}} + \hat{d}\,\frac{d\hat{d}^{\mathbf{T}}}{dc_{\iota}} = 0 \qquad \text{for } l = 0,1,\ \ldots,\ (M-1)/2.\ (A.10)$$

Since the MLE of ω converges almost surely to the true parameter value ω, therefore $(\hat{c},\hat{d})$ converges almost surely to (c^{o},d^{o}). From (A.10) it follows that the vector $[d\hat{c}/dc_{\iota}, d\hat{d}/dc_{\iota}]^{\mathbf{T}}$ converges to a vector which is orthogonal to (c^{o},d^{o}). Since $\underset{n\longrightarrow\infty}{\text{Lim}}\ [d\hat{c}/dc_{\iota}, d\hat{d}/dc_{\iota}]^{\mathbf{T}}$ belongs to the null space of $\underset{n\to\infty}{\lim}(1/n)D(\hat{c},\hat{d})$ and the later converges to a positive semi definite matrix with null space spanned by (c^{o},d^{o}), therefore $[d\hat{c}/dc_{\iota}, d\hat{d}/dc_{\iota}]^{\mathbf{T}}$ converges to zero vector almost surely. The same reasoning holds for $[d\hat{c}/dd_{\iota}, d\hat{d}/dd_{\iota}]^{\mathbf{T}}$ for $l = 0,\ 1,\ldots,\ (M-1)/2$. Therefore $\dot{F}$, the convergence matrix, converges to zero almost surely. This implies the asymptotic stability of the proposed algorithm.

To complete the proof we need the following results

$$\underset{n\to\infty}{\lim}\frac{1}{n}D(\hat{c},\hat{d}) = \underset{n\to\infty}{\lim}\frac{1}{n}E(D(c^{o},d^{o})) = D_{o}$$
$$\underset{n\to\infty}{\lim}\frac{1}{n}R_{\iota}(\hat{c},\hat{d}) = \underset{n\to\infty}{\lim}\frac{1}{n}E(R_{\iota}(c^{o},d^{o})) = 0$$
$$\underset{n\to\infty}{\lim}\frac{1}{n}S_{\iota}(\hat{c},\hat{d}) = \underset{n\to\infty}{\lim}\frac{1}{n}E(S_{\iota}(c^{o},d^{o})) = 0$$

$$\text{for } l = 0,\ 1,\ \ldots,\ (M-1)/2 \qquad\qquad (A.11)$$

Here E stands for expectation and D_{o} is a positive semi definite matrix with null space spanned by (c^{o},d^{o}). The proof of (A.11) is very much similar to that of Theorem A.1 of Kundu (1994) so it is omitted.

BIBLIOGRAPHY

Bai, Z.D., Krishnaiah, P.R. and Zhao, L.C.(1986) *On Simultaneous Estimation of the Number of Signals and Frequencies under a Model with Multiple Sinusoids*, Tech. Report, 86-37, Center for Multivariate Analysis, University of Pittsburgh.

Bai, Z.D., Rao, C.R.and Chow, M.(1989) *An Algorithm for Efficient Estimation of Superimposed Exponential Signals*. Proceedings Tencon, 4th. IEEE Region 10, International Conf., 342-347.

Barrodale, I. and Olesky, D.D.(1981) *Exponential Approximation Using Prony's Method*, The Numerical Solution of Nonlinear Problems (Baker, C.T.H. and Phillips, C., eds.), pp.258-269.

Bresler, Y. and Macovski, A. (1986) *Exact Maximum Likelihood Parameter Estimation of Superimposed Exponential Signals in Noise*, IEEE Tran. Acous., Speech and Signal Processing, ASSP 34 no.5, 1081-1089.

Crowder, M. (1984) *On Constrained Maximum Likelihood Estimation with Non I.I.D. observations*, Ann.Inst. Stat. Math., 36A, 239-249.

Froberg, C.E.(1969) *Introduction to Numerical Analysis*, 2nd Ed., Addison-Wesley, Reading, Mass.

Hannan, E.J.(1971) *Non-Linear Time Series Regression*, Jour. of App. Prob, 8, 767 - 780.

IMSL (1984) *User Manual*.

Kahn. M.H., Mackisack, M.S., Osborne, M.R. and Smyth, G.K. (1992) *On the Consistency of Prony's Method and Related Algorithms*, Jour.Comp. Graph. Stat. 1, 4, 329-349.

Kay, S. M.(1988) *Modern Spectral Estimation: Theory and Applications*, Englewood Cliffs, NJ, Prentice-Hall.

Kay, S.M. and Marple, L. (1981) *Spectrum Analysis - A Modern Perspective*, Proc.IEEE 69 no. 11,1380-1419.

Kumar, R.(1985) *A Fast Algorithm for Solving Toeplitz System of Equations*, IEEE Trans., Acoust. Speech and Signal Proc., ASSP 33, 254-267.

Kumaresan, R. and Scharf, L.L. and Shaw, A.K. (1986) *An Algorithm for Pole Zero Modeling and Spectral Analysis*, IEEE Trans. Acous., Speech and Signal Proc. ASSP 34 no.3, 637-640.

Kumaresan, R. and Tufts, D.W.(1982) *Estimating the Parameters of Exponentially Damped Sinusoids and Pole-Zero Modelling in Noise*, IEEE Trans, Acoust. Speech and Signal Processing, ASSP 30, 833-840.

Kundu, D.(1989) *Some Results in Estimating the Parameters of Exponential Signals in Noise*,Ph.D. thesis, The Pennsylvania State University

Kundu, D.(1993a) *Estimating the Parameters of Undamped Exponential Signals*. Technometrics, 35,2, 215-218.

Kundu, D. (1993b) *Asymptotic Theory of a Particular Non-Linear Regression Model*, Statistics and Probability Letters, 17, 1, 13-17.

Kundu, D. (1994) *A Modified Prony Algorithm for Damped or Undamped Exponential Signals*. Sankhya, Ser. A, 56, 3, 524-544.

Marple, L.(1979)*Spectral Line Analysis by Pisarenko and*

Prony Methods, Proceedings IEEE Conf.on Acoustics, Speech and Signal Processing, 159-161.

Marple, L.(1987) *Digital Spectral Analysis with Applications*, Prentice Hall Signal Processing Series.

Mitra, A. and Kundu, D. (1993) *Asymptotic Properties of the Least Squares Estimators of Superimposed Exponential Signals*, P.C. Mahalanobis Birth Centenary Volume, Ed. A.K. Datta, Indian Statistical Institute, Calcutta, 663-670.

Osborne, M.R.(1975)*Some Special Nonlinear Least Squares Problems*, SIAM Jour. Num. Anal. 12 no. 4, 572-592.

Osborne, M.R. and Smyth, G.K. (1991) *A Modified Prony Algorithm for Fitting Functions Defined by Difference Equations*, SIAM Jour. Sc. Stat. Comput, 12, 362-382.

Osborne, M.R. and Smyth, G.K.(1994) *A Modified Prony Algorithm for Fitting Sums of Exponential Function* To appear in SIAM Jour. Sc. Stat. Comput.

Pisarenko, V.F. (1972) *On the Estimation of Spectra by Means of Non-Linear Functions of the Covariance Matrix*, Geophy. J. Roy. Astron. Soc., Vol 28, 511-531.

Prony, R.(1795)*Essai Experimentale et Analytique*, Jour. Ecole Polytechnique, 24-76.

Rao, C.R.(1988) *Some Recent Results in Signal Processing*, Statistical Decision Theory and Related Topics IV (Gupta, S.S and Berger, J.O., eds.), vol. 2, pp. 319-332.

Rao, C.R. and Zhao, L.C. (1993) *Asymptotic Behavior of Maximum Likelihood Estimates of Superimposed Exponential Signals*, IEEE Trans, Acoust., Speech and Signal Processing, ASSP 41, 1461-1463

Rice, J.A. and Rosenblatt, M.(1988) *On Frequency Estimation*, Biometrika, 75, 3, 477 - 484.

Smyth, G.K. (1985) *Coupled and Separable Iterations in Non-Linear Estimations*. Ph. D. Thesis, Australian National University, Canberra.

Tufts, D.W. and Kumaresan, R. (1982) *Estimation of Frequencies of Multiple Sinusoids: Making Linear Predictions Perform like Maximum Likelihood.*, Proc., IEEE, Special Issue on Spectral Estimation 70,975-989.

Walker, A.M. (1971) *On the Estimation of a Harmonic Component in Time Series with Stationary Independent Residuals*, Biometrika, 58, 21-36.